March 10, 2003 NLDN FLASH AND STROKE DENSITY DATA BASE : NLDNFD

Contents

1 Overview of the Database

NLDN raw data are stored in large ascii files of hundred of thousand of lines. As they are stored on a remote machine, it is required fetching before processing them and as they are larges, all these steps take a while. Thus, a derived database (NLDNFD) was made to fulfill the following criteria :

- Low amount of memory space to be stored on a local machine
- Quick accessibility
- Gather hundred of thousand of data a day in concise and meaningful variables

This data base is made in order to compare the stroke and flash features sensed by the National Lightning Detection Network (NLDN) to cloud-top-temperatures retrieved from the GOES-8 satellite. Thus, the flash and strokes density are gridded in the GOES-8 grid. This database consist of a file a day containing the +CG, -CG, IC, and flash densities computed over 15-minute epochs spawning 3 hours around lgu events from April to September 1998 and from May to October in 1999.

2 What is inside NLDNFD files?

Each '.lga' file (NLDN raw data) gives a timestamp, geolocation, classification in C (cloud) or G (ground), and the peak current value for strokes detected by NLDN. Whatever the peak current values, strokes with the C flag are tallied in the IC densities, strokes flagged G and associated with negative peak currents are included in the -CG densities. Flashes are defined as the ensemble of lightning strokes that borrow a portion of a same pre-ionized channel. Two databases are available, one stored in the directory $***S_PUBLIC/DATABASE/NLDNFD1*** and the second in the directory $***S_PUBLIC/DATABASE/NLDNFD2***. "$***S_PUBLIC***" is currently ***/n/toaster/u/gmolinie/PUBLIC***. Each of theses databases include the -CG, IC and flashes densities defined above. Moreover,
$***S_PUBLIC/DATABASE/NLDNFD1*** include the +CG density defined as G-flagged strokes with positive peak currents while
$***S_PUBLIC/DATABASE/NLDNFD2*** include the +CG density defined as G-flagged strokes with a peak currents greater than 10 kA.

2.1 The Flash Category

The first step is to constitute the flash category. The recipe is given hereafter : each stroke is considered as a potential first return stroke. For each potential first return stroke, the following 150 strokes in the NLDN database are considered as potential subsequent return strokes. They are checked and classified as subsequent return stroke of the current first return stroke if they satisfy the following criteria :

- closer in time to the first return stroke than 300 ms.
- closer in space to the first return stroke than 80 km

If a stroke is selected as a subsequent return stroke, it is no more nor a potential first return stroke neither a subsequent return stroke of an other first return stroke .

2.2 Data Discretization and Storage

Next, the +CG, -CG, IC and flash densities are computed every 15-min epochs each day in the GOES-8 discretization grid.

The grid : Three kinds of GOES-8 data are available at LANL :

- Brightness temperatures derived form the IR channel (4) of GOES-8 encapsulated in JPEG files before date-switch (date-switch defined in the GOES-8 database documentation). Data are interpolated in a fix grid stored on the scheme in the file :
  "$***S_PUBLIC***/GOES/JPEG/pixel_geolocatio_sat_ir_images_beforedate-switch.bin.xdr"

- Brightness temperature derived form the IR channel (4) of GOES-8 encapsulated in JPEG files after date-switch. Data are interpolated in a fix grid stored on the scheme in the files :
  "$***S_PUBLIC***/GOES/JPEG/pixel_geolocatio_sat_ir_images_lat2.txt" and
  "$***S_PUBLIC***/GOES/JPEG/pixel_geolocatio_sat_ir_images_lon2.txt".

- Infrared and visible refractance scaled on 256 bits and stored in .tiff files. Geolocations of GOES-8 pixels are inferred from daily navigation files.

To choose the grid in which the NLDNFD database is computed, the following rules are applied:

- For data of 1999, the pixel geolocation of the grid used is derived of the GOES-8 navigation file of the day for which the densities are computed or if it doesn't exist of the closest in time GOES-8 navigation file.

- For data of 1998, if a navigation file exists, it is used to compute NLDNFD, else the grid is based on the JPEG file grid.

Flash and stroke densities : For each 15-minute period of 3-hour epochs around FORTE overpasses the North America (NA) (FORTE above the horizon regarding the NA and armed), the number of flashes and strokes of each categories are counted in each mesh of the discretization grid. Next, a compression scheme is applied. In each grid mesh, the maximum number of lightning among all the categories is determined, divided by the grid mesh area, and stored as 32 bit floating point data. It is a normalization factor (Fdn). Next, the flash or stroke densities for all the lightning categories are computed, divided by Fdn multiplied by 32767 and stored as integer. If Fd is the stroke density for any given category, the previous kitchened can be written :

$$Fdn = MAX(Fd); Fd = FIX(\frac{Fd}{Fdn} \times 32767) \quad (1)$$

That insure a good accuracy even if the flash or stroke density is very low regarding Fdn. For example, if $Fd = 10^{-3} \times Fdn$, the uncertainty on the stored flash density is less 2.5%. Only, the grid indexes and the values of the flash and stroke densities where at least one of the flash or stroke densities are non null, are stored.

A NLDNFD daily file weights less than 1 MO of memory and the complete database (6 months in 1998 and 6 months in 1999) less than 200 MO.

3 Tools to Handle the Database

3.1 Plot Flash or Stroke Densities

IDL programs useful to read and to plot the database are available in :
"$***S_PUBLIC***/NLDNFD "
The main routine is : read_grid_flash.pro. The calling sequence is :
READ_GRID_FLASH, Efile,Syst=Sdevice $
,First_scene=Pscene_min, $
Last_scene=Pscene_max, Zoombox=Fzoombox,Goes=SGoes ,Sconcat=Sconcat

Efile is a daily file containing the flash and stroke densities. This files are available on the scheme at :
"$***S_PUBLIC***/DATABASE/NLDNFD? ". ? being 1 or 2
Their names are of the form YYYYMMDD.nldnfd. The keyword "Syst" determine the device in which the graphic outputs are generated. "X", "CGM" and "PS" devices are available. "First_scene" corresponds to the index of the first scene to plot (default value is 1) and "Last_scene" to the last scene to plot (default value is the last scene included in Efile). Zoombox indicates the longitudes and latitudes defining a region to zoom in and if "Sgoes" is set a map of GOES-8 brightness temperature is also plotted. If Sconcat is set and the device is 'ps' or 'cgm' the output files are compressed and archived as explained in the example 2. In figure 1, an example of plot is given corresponding to the calling sequence :

READ_GRID_FLASH,'/nh/toaster/u/gmolinie/IDL_EXE/NLDN/NLDNFD/19990802.nldnfd', $
Syst='ps',First_scene=23, Last_scene=24,Zoombox=[-88,23,-76.,32],/Goes

A second example is given in figure 3, it concerns data whose the grid is consistent with a jpeg file. The calling sequence as well as the answers of *idl* are shown in figure 2. When a message is printed by routines from $***S_PUBLIC***, it begins by the upper-case name of the routine. Running **READ_GRID_FLASH** creates 8 postscript file (19980616.nldnfd.1.gic.ps, 19980616.nldnfd.1.ic.ps, 19980616.nldnfd.1.gncg.ps, 19980616.nldnfd.1.ncg.ps, 19980616.nldnfd.1.gpcg.ps, 19980616.nldnfd.1.pcg.ps, 19980616.nldnfd.1.flash.ps, 19980616.nldnfd.1.gflash.ps). If 'Sconcat' is set and the device is 'ps' or 'cgm' the output files are compressed and archived. If the device is for instance 'ps', 8 encapsulated poscript files (19980616.nldnfd.1.gic.epsi, 19980616.nldnfd.1.ic.epsi, 19980616.nldnfd.1.gncg.epsi, 19980616.nldnfd.1.ncg.epsi, 19980616.nldnfd.1.gpcg.epsi, 19980616.nldnfd.1.pcg.epsi, 19980616.nldnfd.1.flash.epsi, 19980616.nldnfd.1.gflash.epsi) are created via a shell script (launching "***ps2epsi*** that induce error messages). Their name "YYYYMMDD.nldnfd.FDKIND.tgz" include the date of the NLDNFD file (YYYYMMDD), followed by 'nldnfd' and then the scene number followed by the name of the plotted parameter (FDKIND equal *flash*, *ic*, *ncg*, *pcg gflash*, *gic*, *gncg*, *gpcg*). When FDKIND begins with the letter "g", the file consists to the map of the brightness temperature on which crosses indicate the location of pixels where the given kind of stroke or flash densities are non-zero. The color bar on the right of the map indicates the brightness temperature (°C). The other map, for which FDKIND doesn't begin with a "g", indicate stroke or flash densities ($km^{-2}(15min)^{-1}$). Graphs of *ic*, *ncg* and *pcg* are stroke densities and, only the graph named *flash* displays flash densities. The shell script also archives and compress the ".epsi" files in a .tgz file in order to save memory space. The name of the archive looks like "YYYYMMDD.nldnfd.epsi.tgz", that is, in the current example : *"19980616.nldnfd.epsi.tgz"*
The shell command ***tar tvfz 19980616.nldnfd.epsi.tgz*** provide a list of the archived files.
The shell command ***tar xvfz 19980616.nldnfd.epsi.tgz*** extract all the archived files.
The shell command ***tar xvfz 19980616.nldnfd.epsi.tgz 19980616.nldnfd.1.pcg.epsi*** extrat only the file: ***"19980616.nldnfd.1.pcg.epsi"***
Moreover, in order to have a quick look of the plotted ".epsi" file, jpeg files are provided that can be seen and even animated by any image processing software. On LINUX systems where the software 'ee' is available, the command "ee 19980616.nldnfd.*.ncg.jpg" animate the negative CG density maps.

3.2 Read the NLDNFD files

Two kinds of informations are available in the NLDNFD files: the discretization grid and, the flash and stroke densities. One routine allows to retrieve both of these parameters (retrieve_grid_flash.pro), two others allow to retrieve either the grid (read_grid.pro) or, the flash or stroke densities (read_flash.pro).

3.2.1 Retrieve both the NLDNFD grid and data

$***S_PUBLIC***/RETRIEVE_GRID_FLASH allows to retrieve the grid and, the flash and stroke densities for a number of scenes limited to 20 because a higher number of scenes involves a too large memory allocation,.

```
    PRO RETRIEVE_GRID_FLASH, Efile ,STgrid2D, STf_d2D $
,First_scene=Pscene_min $
,Last_scene=Pscene_max

    ; Purpose : Read the grid and flash density in the NLDNFD files and
; set arrays containing the values. One set of
; lightning densites are provided for each scene.

    ; Inputs : Efile : path and name of the file to read

    ;Outputs : STgrid2D: provides the latitude and longitude of the
; pixels. The grid is stored in 2D
; struture containing 2 elements (Flon:
; longitude and Flat: latitude)
;
; STf_d2D: provide the flash, -CG, +CG, IC under the
; structure elements: Fflash, Fncg, Fpcg, Fic,.
; STf_d2D is 3D structure. The first 2
; dimensions are identical to these of
; STgrid2D. Thus, STgrid2D[i,j] provides the
; geolocation of the center of the pixel
; whose STf_d2D[i,j,*] provides the
; lightning densities.
;
; Keyword :
; First_scene: first scene to plot, if not set
; First_scene = 1
; Last_scene : last scene to plot, if not set
; Last_scene = the last scene available in Efile
;
;
```

3.2.2 Retrieve the NLDNFD grid

- To read the NLDNFD grid, one must use the program :
  $***S_PUBLIC***/NLDNFD/read_grid.pro

  Calling sequence : READ_GRID,Iur,STsize,STgrid2D, Ifile_size=Nnldnfs, $
  Ifile_done=Nnldnfsd, Ifile_error=Nfile_error, Sday=Sday

Input :
Iur : logical unit for the nldndf file

Outputs :
STsize : grid size (STsize = { ix:0 , iy:0})
Stgrid2D : lat, lon of each grid meshes (STgrid = REPLICATE({Blon : BYTARR(STsize.iy) , Blat : BYTARR(STsize.iy)}, STsize.ix)

Keywords :
Ifile_size : size of the file to read
Ifile_done : size of the file already read
Ifile_error : 1 if the end of the file pointed by Iur is reached before expected
Sday : day to treat (This keyword is required)

Retrieve the file used to define the NLDNFD grid In this goal, use the function:

FUNCTION RETRIEVE_GOES_FILE,Iyear, Imonth, Iday
;+
; Purpose : Return the name of the GOES-8 file whose the grid was used
; to compute the NLDNFD grid. In case of a jpeg file
; return the closest to noon available jpeg file.
; The rules applied in the routine
; flash_density.pro to define the grid are :
; the grid is based on the .nav file if exist
; else the grid is based on the jpeg-file grid if one exist in
; the day.
; If none of the .nav and jpeg file exist, the grid
; is based on the default .nav file (19991128__G08I04.nav)
;
;
; Inputs : Iyear : year of the NLDNFD file
; Imonth : month of the NLDNFD file
; Iday : day of the NLDNFD file

3.2.3 Retrieve the NLDNFD data

- To read the NLDNFD data, one must use the program :
  $***S_PUBLIC***"/NLDNFD/read_flash.pro"
  Calling sequence : READ_FLASH,Iur,STsize,Sdate,Ncount,Nindex,STf_d1D,STf_d2D , $
  Ifile_size=Nnldnfs,Ifile_done=Nnldnfsd,Ifile_error=Nfile_error

  Inputs :
  Iur : logical unit of the nldnfd file to treat
  Stsize : grid size (STsize = { ix:0 , iy:0})

  Outputs :
  Sdate : date of the readed scene (Sdate = 'YYYYMMDD1_HHMMSS')
  Ncount : # of grid meshes where Flash or +CG or -CG or ID are non null
  STf_d1D : 1D array structures containing the flash, +CG, -CG and ID densities (STf_d1D =

REPLICATE({Ipcg: 0 , Incg: 0, Iic: 0, Iflash: 0, Fn: 0.},LONG(STsize.ix)*LONG(STsize.iy)))

STf_d2D : 2D array structures containing the flash, +CG, -CG and ID densities (STf_d2D = REPLICATE({ Fpcg: 0. , Fncg: 0., Fic: 0., Fflash: 0.}, STsize.ix,STsize.iy))

==========

Gilles Molinie
NIS-1 Group Office (Space and Atmospheric Sciences)
TA-3/1888/01U, Bikini Atoll Rd., SM-30
Los Alamos National Laboratory
P. O. Box 1663, MS D466, Los Alamos, NM 87545

e-mail : gmolinie@lanl.gov Phone : 505 667 6104

==========

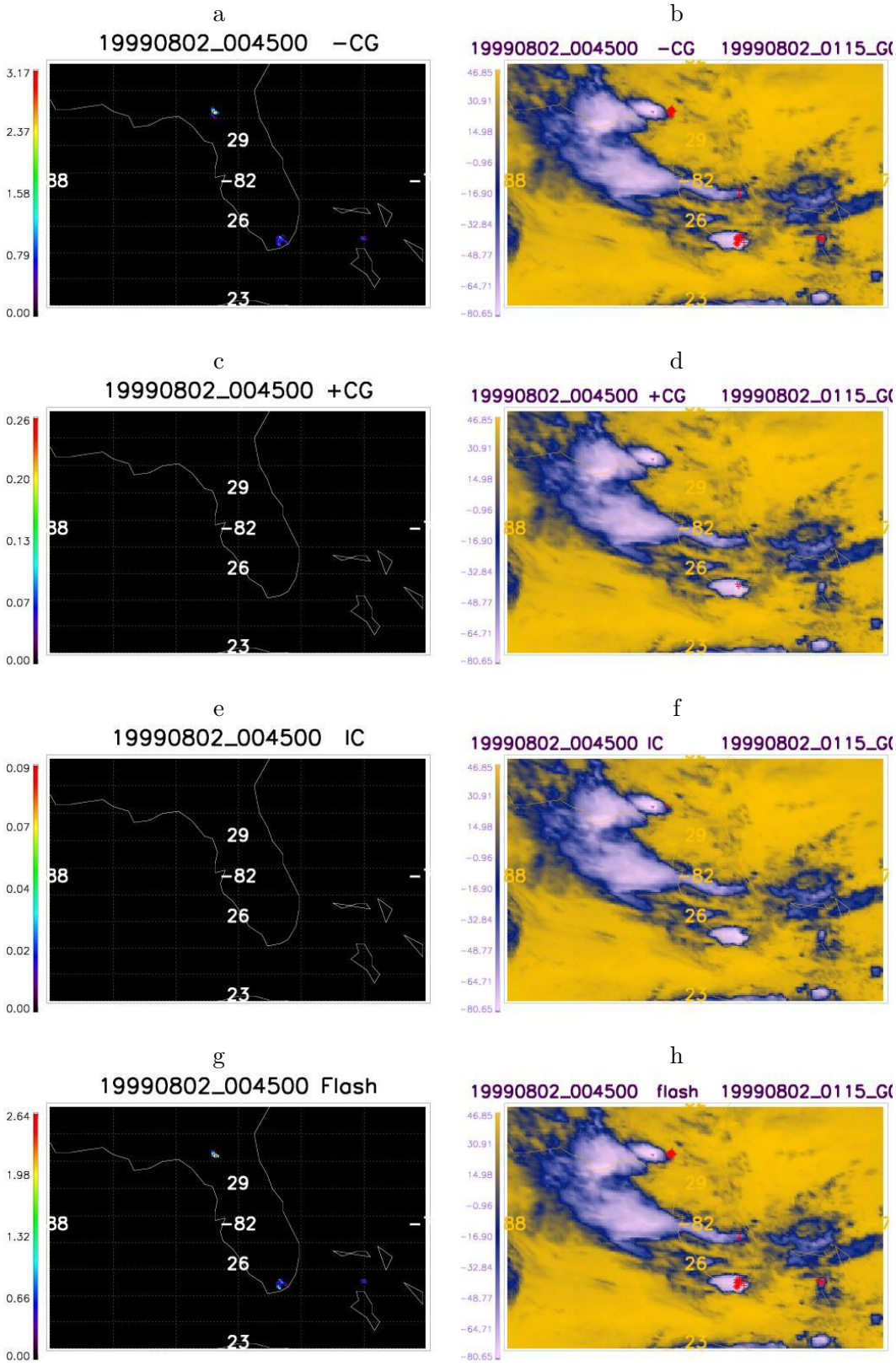

Figure 1: a : -CG stroke densities the August 2, 1999 at 00:30 (the colorbar indicating the stroke density in *stroke* km^{-2} *hour*$^{-2}$); b : GOES-8 brightness temperature (scale indicated by the color bar) on which are superimposed the location of non-zero -CG stroke density pixels; c & d same as a & b but for +CG; e & f same as a & b but for for IC; g & h same as a & b but for flash.

```
IDL> read_grid_flash,'/nh/toaster/u/gmolinie/IDL_EXE/NLDN/NLDNFD1/19980616.nldnfd

READ_GRID_FLASH : Read the NLDNFD grid

READ_GRID  : Rebuild the grid in a 2D array
READ_JPEG : read file/n/projects/forte/tmp/sat_ir_9806160022.jpg
READ_GRID : NLDNFD grid based on grid discribed in pixel_geoloc_sat_ir_im_beforec

READ_GRID_FLASH : Prepare to plot NLDNFD
% LOADCT: Loading table Rainbow + white
READ_GRID_FLASH : Device = PS
READ_GRID_FLASH : Read the NLDNFD data

       1 Scenes done
RETRIEVE_GOES_DATA:  RUN JPEG_2_TBRIGHT

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.flash.ps

MAP_FD_GOES : For plotting purpose, the array resolutions are divided by
       2
% LOADCT: Loading table Haze
% LOADCT: Loading table Rainbow + white
% LOADCT: Loading table Haze

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.pcg.ps

MAP_FD_GOES : For plotting purpose, the array resolutions are divided by
       2
% LOADCT: Loading table Haze
% LOADCT: Loading table Rainbow + white
% LOADCT: Loading table Haze

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.ncg.ps

MAP_FD_GOES : For plotting purpose, the array resolutions are divided by
       2
% LOADCT: Loading table Haze
% LOADCT: Loading table Rainbow + white
% LOADCT: Loading table Haze

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.ic.ps

MAP_FD_GOES : For plotting purpose, the array resolutions are divided by
       2
% LOADCT: Loading table Haze
% LOADCT: Loading table Rainbow + white
% LOADCT: Loading table Haze

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.pcg.ps

PLOT_MAP : For plotting purpose, the array resolutions are divided by        2

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.1.ncg.ps

PLOT_MAP : For plotting purpose, the array resolutions are divided by        2
       2 Scenes done
RETRIEVE_GOES_DATA:  RUN JPEG_2_TBRIGHT

PLOT  FILE :  /nh/toaster/u/gmolinie/TEMP/19980616.nldnfd.2.flash.ps

MAP_FD_GOES : For plotting purpose, the array resolutions are divided by
       2
% LOADCT: Loading table Haze
% LOADCT: Loading table Rainbow + white
% LOADCT: Loading table Haze
```

Figure 2: Calling sequence and idl dialogs relative to the figure 3.

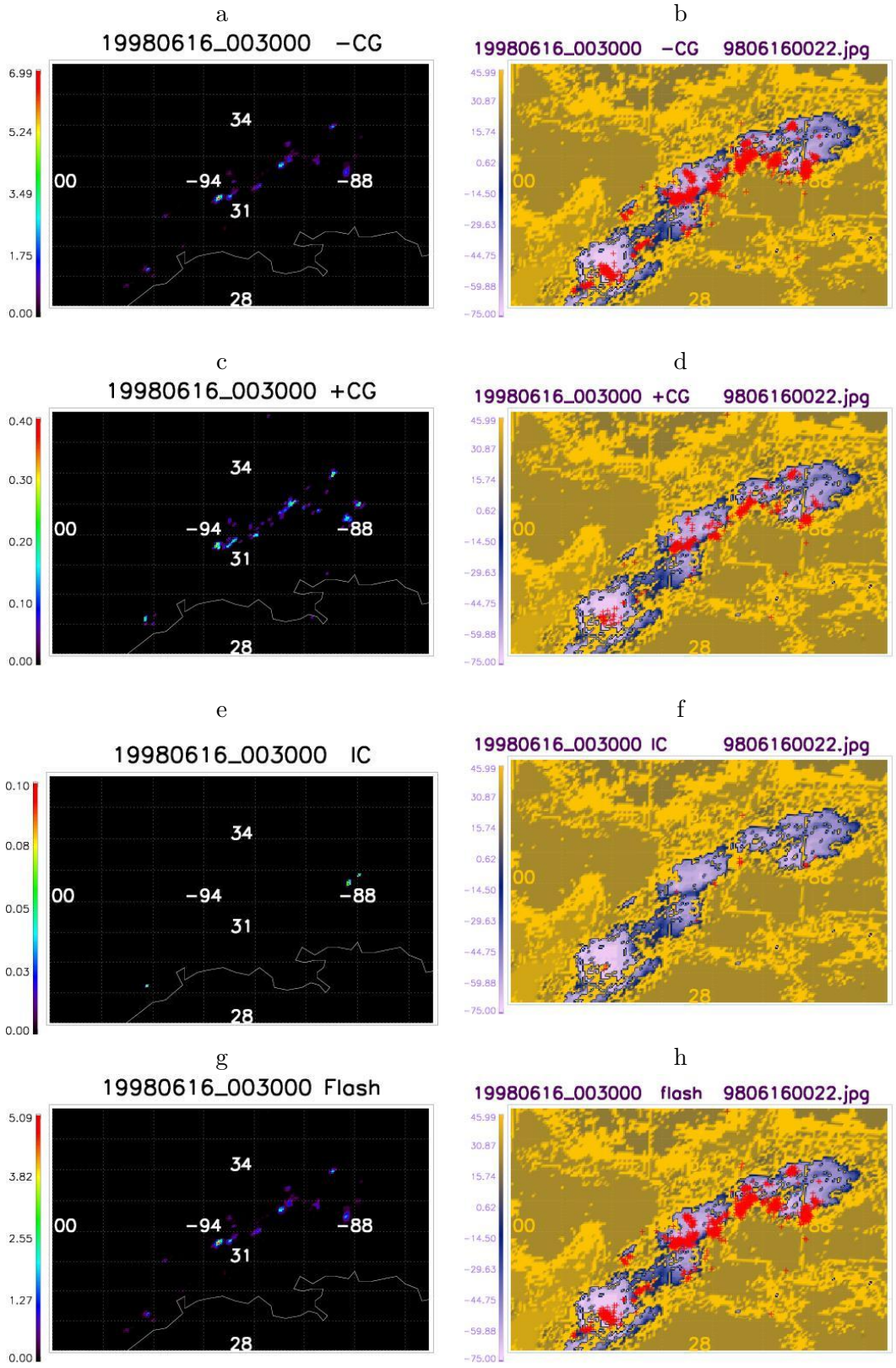

Figure 3: a : -CG stroke densities the June 6, 1998 at 00:30 (the colorbar indicating the stroke density in *stroke* km^{-2} *hour*$^{-2}$); b : GOES-8 brightness temperature (scale indicated by the color bar) on which are superimposed the location of non-zero -CG stroke density pixels; c & d same as a & b but for +CG; e & f same as a & b but for for IC; g & h same as a & b but for flash.